

*11/8
03-d-d*

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/508,821A DATE: 02/23/2001
 TIME: 16:03:53

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02232001\I508821A.raw

2 <110> APPLICANT: ROULEAU, Guy A.
 3 JOOBER, Ridha
 4 BENKELFAT, Chawki
 6 <120> TITLE OF INVENTION: POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF
 8 <130> FILE REFERENCE: 12948-10
 10 <140> CURRENT APPLICATION NUMBER: 09/508,821A
 11 <141> CURRENT FILING DATE: 2000-05-26
 13 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00884
 14 <151> PRIOR FILING DATE: 1998-09-18
 16 <150> PRIOR APPLICATION NUMBER: 2,216,057 CA
 17 <151> PRIOR FILING DATE: 1997-09-19
 19 <160> NUMBER OF SEQ ID NOS: 6
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 328
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Mus musculus
 28 <400> SEQUENCE: 1
 29 agggcagcca ctttccccagg cattcccaagt ctctccactac tatgccccaa 60
 30 cagtgcaggg tgggtggcagg gggggccact octacaagag ctgcacagca ccatctgccc 120
 31 agccctcatga taggcgcattc agtgccatcg cgaacctggc tcacgggcaa cgggtccaga 180
 32 atcttcacgc ttaccagcc tggccgccttg gctacagcaga gcacgcacaa gcaacttcaag 240
 33 gccgtcacca caccaggaa acactccact accagaacct cgccaaatgc caaacatctatg 300
 34 gacagcaagg ccagggtctac tgccacc 328
 37 <210> SEQ ID NO: 2
 38 <211> LENGTH: 330
 39 <212> TYPE: DNA
 40 <213> ORGANISM: Homo sapiens
 42 <220> FEATURE:
 43 <221> NAME/KEY: modified_base
 44 <222> LOCATION: (59)
 45 <223> OTHER INFORMATION: N = A or G or C or T/U
 47 <400> SEQUENCE: 2
 W--> 48 tccttccca ctcctccac ctactctcc tctgtcccaag ggtgggtggc agggggccna 60
 49 ctccataaag aatgcacag caccgactgc ccagcccccatt gacaggccgc tgactgccag 120
 50 ctccagccatg gccccggggc aeggggtccaa gaatcttcat gcttaccatg cggggccct 180
 51 cagctatgac cagcagcagc aacgcacagca gcacgcacgc cagcagcagc aageccctca 240
 52 gageeggac catgccccagg aaacccttca ttacaaaaac ctcgcacatg atcagcacta 300
 53 cgggcagcaa ggccagggtct actggccagcc 330
 56 <210> SEQ ID NO: 3
 57 <211> LENGTH: 23
 58 <212> TYPE: DNA
 59 <213> ORGANISM: Artificial Sequence
 61 <220> FEATURE:
 62 <223> OTHER INFORMATION: Description of Artificial Sequence:
 63 oligonucleotide
 65 <400> SEQUENCE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:53

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\1508821A.raw

66 ggccggcagg gtccagaatc ttc
69 <210> SEQ ID NO: 4
70 <211> LENGTH: 23
71 <212> TYPE: DNA
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: Description of Artificial Sequence:
76 oligonucleotide
78 <400> SEQUENCE: 4
79 agcactacgg gcagcaaggc cag
82 <210> SEQ ID NO: 5
83 <211> LENGTH: 6022
84 <212> TYPE: DNA
85 <213> ORGANISM: Homo sapiens
87 <400> SEQUENCE: 5
88 ggatccaggca ggcggcaagg gatgagggg cgaaaaattgc tctgtctaaat gcttttgagc 60
89 tgcaggaaag ggctggggat gatgggtggg ggacattggg gaggagctgg caatggcgq 120
90 ggggggggggg ggtatgcctcc cctatggcaag ctatggatgc ggcgggttttgc cttcccgat 180
91 cttatggccgtt cttatggcaag actcgtgc agggatgtg ggatggggaa ttacagagca 240
92 caccctccgt acacagaatg tttatggatgc cttatggcaag ctatggatgc tgatggggggc tcaggcgaaq 300
93 ggggggactat taatggatgc gggggggggc agggatgtg ggatggggggc ttggatgggtgg 360
94 cttttctgggg cttatggatgc gttatggatgc gttatggatgc tttatggatgc tttatggatgc 420
95 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 480
96 ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 540
97 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 600
98 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 660
99 agctatgggg gttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 720
100 tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 780
101 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 840
102 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 900
103 tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 960
104 tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1020
105 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1080
106 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1140
107 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1200
108 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1260
109 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1320
110 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1380
111 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1440
112 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1500
113 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1560
114 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1620
115 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1680
116 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1740
117 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1800
118 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1860
119 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1920
120 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 1980
121 cttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc tttatggatgc 2040

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:53

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

```

122 ctgagcggtt ccgaggaccc actggagcgc agcttcctct actgcaaccca ggccccggc 2100
123 agccctggca gggtaaacag caactcgaaag gcacaaggccc aglcccgltc caccgtttct 2160
124 gtgacctetc ctgacgacat gtccacccaa tctgacgact ctttccaaag cctcacacggc 2220
125 agtctggccg ttcacgactt ctcacggctt gttggcggtt agggggactg tccggggctg 2280
126 ctgctcggcg ccctggcata ggaggacccg gcttcggaga tcttggggct gcaaggaaagcc 2340
127 atcggtgaga aggccgacaa agcttggctt gaaagcaceca gcttggctca ggacacggc 2400
128 aagecaccct ttcctgttggaa gaaaccacage gcttgcetgg actctgttgc caaagagtgc 2460
129 tggcccccggc ctggggagcc ggagggccctg cccqactctt tgcagcttggaa caaggggggc 2520
130 aatgecaagg acttcacggc agggcttggt gaaagaccctt ccttggccctt cgctacgct 2580
131 gaccccaaaa agacaacttgg ttccttcctt ttgggttacca agccacccctt tgggggtctt 2640
132 gtcacggacc caactacacg agtttttgc ttttccccc acacaacccg tggcagctca 2700
133 gccggacccg ccaacccctt tgcgtggccaa qqqqaaaaacc tgggggatgc ttgtccccagg 2760
134 tggggatttc accctggcggc gcttaccaag ggcttggcgc aggggtggaa ggccttcagat 2820
135 ggcacfcagca aaggggacac ccattggatc tccgttgc tggcttccaa ggaggaggac 2880
136 ccccttgggg agaaagggttgc ctcgttggcc 9999ac1tca acaaggaggaa ggllgggtgg 2940
137 gtgaaggagg aggccgggttgc gcttgcgttgc tggccgggggg tggccaaaggc tgaccgggttgc 3000
138 ctggggacca gggggccactg ctgttcaccc gccgacttcg gggacccccc actctgttgc 3060
139 cccacccggc gggaggaggaa ctcgttggactt gaggaggatc actcttcctt atgttgc 3120
140 ctggggccggc cccggggccatc gcttggcatc caggggccccc tgcacccca gggcccaactc 3180
141 atctgcacca aggaggaggat gggggggctt ctggggccatc aggccggctt gggcccttcg 3240
142 tggcaccctt caggggggatc ctgttgcactt ctggggccatc cagtggggccatc cggatcaaaag 3300
143 gtcacggatc ggttgttgc ttccttgcata cacatgttgc aagggttgc aagggttgc 3360
144 gggggccggc ttcacggggaa ttccacccacc tcggggccctt ctgttggccaa yaagcccaac 3420
145 aagectgttgc tggccggggcc gcccategc aagaaaggac ctgttgcacccy gggcaaaagc 3480
146 ttacggggcc ctgggggttgc cccggggctt cccggggccgg aggacttccctt atgcggggca 3540
147 ccagtgttgc ccaaaagactt ctgttgcctt gaaatcttgc aaggggcccccc ccaggggacac 3600
148 atggaaagggg ctggggccccc aggccgggggg gcttggggaa ggttcccccgg gatgttgcact 3660
149 ctgttgcata cggcccttgc tggggccggc acggccggac ccccaaggcc taccacccacc 3720
150 cctgttgcata cggccatccact gggggccggc cccgttgcaccc gggcaaggccg 3780
151 gtggggaaatc ctcaccccaaa ggttgcetttccaa agccccccggc cctgttgccttgc 3840
152 gcttcggccca gcaacccggat gggccatccaa accaaggaggaa cagacttccatc cagccggccat 3900
153 ggcaaggacc agcgttccatc gatecttggcc tcaacggccca aaaccccgatc gatcttccac 3960
154 tccaaaggcc gggggccctt tggggccggc ctggggccatc ggggttgcaccc caaaggatc 4020
155 ctgcacccaa cccatcttgc cccacccatc aagggttgcata cccggccatccaa gaaggaggcc 4080
156 agggtgatcc acggggccggc gttccacccatc ctgttgcaccc gcaaggatc ctgttgcaccc 4140
157 ccccttccatc ctgttgcataaagg gaaatgttgc ttcacccatc cgggttccatc caaggatc 4200
158 aacctgttgc tggggccatc cccggccatc aaggggccatc ctgttgcataat gggggggatc 4260
159 ggaaggaggaa gggccatc ggttcccccacttccatc gggatgttgc ttcacccatc 4320
160 aageccaaaggcc cccatccatc aatggccggc ctgttgcataaa gtcacccatccaa cccggggccatc 4380
161 ccatcttgc ctcaacgttgc ctgttgcata ggttccatc gggccatccaa gaccaaggatc 4440
162 ctggccacccccc ggaaggggccg gggccatc gttccatccaa tgggttgcata gttccatccatc 4500
163 cccatccatc agaaggatc gttccatccaa aaggggccatc ctgttgcataat ttcacccatc 4560
164 ccatcttgcatc cccatccatc aatggccggc aaggggccatc ctgttgcataat gggggggatc 4620
165 gaagaaggcc tggtaaatgt gggccatccaa cccatccatc aatggccggc ctgttgcataat gggggggatc 4680
166 ttatgttgcataa atccacccaa cccatccatc aaggggccatc ttcacccatc gttccatccatc 4740
167 ttttctatc actgttgcataa aacccggggcc ttcacccatc cccatccatc gttccatccatc 4800
168 gggacttgc tggccatccaa aaggggccatc cccatccatc gggccatccaa cccatccatc 4860
169 tccaaaggcc ggttgcataa gggccatccatc ctgttgcataat gttccatccatc 4920
170 gacaggccatc gttccatccatc aaggggccatc gttccatccatc gttccatccatc 4980

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:53

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

```

171 ccaaaggatgg aggagctggg ccctgcctcc cagccccccc agggcaggcc ctgccagcc 5040
172 cagaccaaggc cacagaaaca gcccggccac accaactaca gcagtttattc caagcgaaag 5100
173 cgcctcaact cggggcgggc caagaacacc acctcttccac cctgttaaggg gcgtggcaag 5160
174 cgaacgacac aacagcaggcg gtcggccctg gatcccccgag agccgtaaat cccgcctcaag 5220
175 tacattttctt ctgtcaagcg gtcgggtca gacagcccgaa ccccccgtt ctcacccctt 5280
176 gtgcgggtgg agaaggcgaga cgcgttccac accatgtca ctgttgtcaa ctcccttgaa 5340
177 gatgcggcca aacccceacag qaagcccttc tccctctgcct cctcttcctt atcccttgtcc 5400
178 tcgttctctt tggtatgcacq cggggcccttc ctggccacac tccctggagg ctccatctgt 5460
179 cagccggccg ccccttcgtcc cctcttcctt aegatgcact tggggctgtt gggttccaaag 5520
180 qcclgagta ctcttgcct tgggtgtcc ctcttgcctt accccggccaa cttaaggac 5580
181 ctggggacc tctgtggcc ctactaccct gaacacttcc tccccaaaaa qaagccaaaa 5640
182 ctcaaggaga aagtgcggcc agaaggccacc ttgtggggagg ctccgttgcct qcttgagaga 5700
183 acactcaaaag gtcggcgggtg tgcagtttcc gcaacttgcg ggaagcccccc cagggtgacgg 5760
184 cccagcttgc cccggccaaagg agggcccaact ggcgcaccatgt gccccggggcc tggccggag 5820
185 gtcgcagacg tgcatactgtt gtgtatggccgg ggaggatggg ggegaggagg cagcccccc 5880
186 cgacaagggtt cggcaaacatq agtcagccat gggaggctccg gcaagcccccc qgqqqqqagqc 5940
187 ccaggcggcac tgggtgcattt aggcctgttc cgtgtggacc gggggctt accttgtggc 6000
188 cggggaaactt tttgggtgtcc ag 6022
191 <210> SEQ ID NO: 6
192 <211> LENGTH: 22
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: Description of Artificial Sequence:
197 oligonucleotide
200 <400> SEQUENCE: 6
201 tggccttgtct gcccgttagtg ct

```

22

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/508,821A DATE: 02/23/2001
TIME: 16:03:54

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

L:48 M:341 W: (46) "n" OR "Xaa" used, for SEQ ID#:2